LIVER

Diabetes increases the risk of hepatocellular carcinoma in the United States: a population based case control study

J A Davila, R O Morgan, Y Shaib, K A McGlynn, H B El-Serag

....

See end of article for authors' affiliations

....

Correspondence to:
Dr H B El-Serag, Houston Veterans Affairs Medical Center, 2002 Holcombe Blvd. (152), Houston, TX 77030, USA; hasheme@ bcm.tmc.edu

Background: Diabetes has been associated with an increased risk of hepatocellular carcinoma (HCC) in studies of referred patients. This is the first population based case control study in the USA to examine this association while adjusting for other major risk factors related to HCC.

Methods: We used the Surveillance Epidemiology and End-Results Program (SEER)-Medicare linked database to identify patients aged 65 years and older diagnosed with HCC and randomly selected non-cancer controls between 1994 and 1999. Only cases and controls with continuous Medicare enrolment for three years prior to the index date were examined. Inpatient and outpatient claims files were searched for diagnostic codes indicative of diabetes, hepatitis C virus (HCV), hepatitis B virus (HBV), alcoholic liver disease, and haemochromatosis. HCC patients without these conditions were categorised as idiopathic. Unadjusted and adjusted odds ratios were calculated in logistic regression analyses.

Results: We identified 2061 HCC patients and 6183 non-cancer controls. Compared with non-cancer controls, patients with HCC were male (66% v 36%) and non-White (34% v 18%). The proportion of HCC patients with diabetes (43%) was significantly greater than non-cancer controls (19%). In multiple logistic regression analyses that adjusted for demographics features and other HCC risk factors (HCV, HBV, alcoholic liver disease, and haemochromatosis), diabetes was associated with a threefold increase in the risk of HCC. In a subset of patients without these major risk factors, the adjusted odds ratio for diabetes declined but remained significant (adjusted odds ratio 2.87 (95% confidence interval 2.49–3.30)). A significant positive interaction between HCV and diabetes was detected (p<0.0001). Similar findings persisted in analyses restricted to diabetes recorded between two and three years prior to HCC diagnosis.

Conclusions: Diabetes is associated with a 2–3-fold increase in the risk of HCC, regardless of the presence of other major HCC risk factors. Findings from this population based study suggest that diabetes is an independent risk factor for HCC.

T he incidence of hepatocellular carcinoma (HCC) has increased significantly over the past two decades in the USA.[1][2] The age adjusted incidence of HCC increased from 1.3 per 100 000 for the period 1981–1983 to 3.0 per 100 000 for the period 1996–1998.[2] The cause of this increase in HCC is incompletely understood. Approximately half of this increase is attributed to hepatitis C virus (HCV), while a minimal or no increase has been related to hepatitis B virus (HBV) or alcoholic liver disease.[3][4] In general, approximately 15–50% of HCC cases remain idiopathic, suggesting that other risk factors are responsible for this increase in HCC.

Diabetes has been suggested as a potential risk factor for HCC. However, the causal association between diabetes and HCC is difficult to study. On the one hand, diabetes is a risk factor for non-alcoholic liver disease (NAFLD), including its most severe form, non-alcoholic steatohepatitis (NASH),[5–10] which can lead to cirrhosis and subsequently HCC. On the other hand, end stage liver disease itself can cause glucose intolerance and overt diabetes.[11] Moreover, some causes of liver disease (for example, HCV and haemochromatosis) have been associated with an increased risk of diabetes. Several studies have examined the association between HCC and diabetes.[12–18] Earlier studies reported no association between diabetes and HCC while more recent studies have identified diabetes as a risk factor for HCC.[12][13][15][17][18] For example, a cohort study conducted in patients identified at hospitals of Veterans Affairs found a greater than twofold increase in the relative risk of HCC among male veterans with diabetes in the absence of HCV, HBV, and alcoholic cirrhosis.[12] However, the generalisability of these results to the overall US population may be limited; all except for one small study[17] were based on

referral samples. Moreover, selection bias may have occurred in which patients with diabetes were more likely to be tested and diagnosed with HCC, and therefore more likely to be enrolled in these studies.

We conducted a population based case control study using the linked records of patients with HCC identified from the Surveillance, Epidemiology, and End-Results (SEER) registries to Medicare claims.[19] The purpose of this study was to determine whether an association exists between HCC and diabetes in Medicare enrolled patients age 65 years and older. Furthermore, we examined the effect of diabetes in relation to the presence of other risk factors, including HCV, HBV, alcoholic liver disease, and haemochromatosis.

METHODS
Data source
The SEER-Medicare database is the linkage of SEER registry information with Medicare claims data. The SEER programme is an ongoing contract supported programme of the National Cancer Institute to collect population based cancer incidence and survival data. Since 1992, the SEER programme has collected data on incident cancer cases from 11 cancer registries in five states (Connecticut, Hawaii, Iowa, New Mexico, and Utah) and six metropolitan areas (Los Angeles, San Francisco/Oakland, San Jose, Detroit, Seattle,

....

Abbreviations: SEER, Surveillance, Epidemiology, and End-Results Program; HCC, hepatocellular carcinoma; HBV, hepatitis B virus; HCV, hepatitis C virus; NAFLD, non-alcoholic fatty liver disease; NASH, non-alcoholic steatohepatitis; HMO, health maintenance organisation; OR, odds ratio; HIV, human immunodeficiency virus

and Atlanta) that account for approximately 14% of the population in the USA.[19] For each case identified, the SEER programme collects demographic features, date of cancer diagnosis, cancer site, and method of diagnosis (histology, cytology, microscopic confirmation (method not specified), laboratory test/marker study, direct visualisation or positive radiology tests). The International Classification of Diseases for Oncology (ICD-O-2) is employed by SEER to classify primary site and histological type for cancer.[20]

Medicare is the primary health insurer for approximately 97% of individuals aged 65 years and older in the USA. Medicare claims data are collected for both Medicare part A and part B benefits. Approximately 95% of Medicare beneficiaries are covered by both part A and part B benefits. Persons less than 65 years of age are eligible for Medicare benefits due to disability or end stage renal disease; these patients were excluded from the study as they are significantly different from patients aged 65 years and older in their clinical characteristics and demographic features.[19] Claims information from inpatient hospitalisations covered by Medicare part A benefits are included in the Medicare Provider Analysis and Review files. This file contains up to 10 diagnosis and 10 procedure codes using ICD-9-CM codes. Medicare claims data for all part B covered benefits include physician/supplier services, and contain ICD-9-CM diagnosis codes and Current Procedural Terminology 4 codes for all billed claims.

The linked SEER-Medicare data is a collaborative effort by the National Cancer Institute, the SEER registries, and Centers for Medicare and Medicaid Services. This database contains Medicare part A and part B claims data beginning in 1991 for all Medicare enrolled patients identified by SEER registries. The linkage captures approximately 93% of patients in the SEER database aged 65 years and older. Additional details regarding this linkage were described elsewhere.[19]

The SEER-Medicare database also contains Medicare files for beneficiaries without cancer residing in the SEER areas. A 5% random sample of beneficiaries is provided to serve as a comparison ''control'' population for Medicare patients identified by the SEER registries. The contents of Medicare files for the ''control'' sample are identical to those available for patients with cancer.[19]

Study population

To include both cases and controls with equal exposure to risk factor information, we selected only those with continuous enrolment in Medicare parts A and B for at least the three years prior to cancer diagnosis or the randomly selected index date for non-cancer controls. We excluded both cases and controls enrolled in a health maintenance organisation (HMO) during this time period as Medicare HMO plans have not been required to submit individual claims to CMS for specific services received by patients enrolled in Medicare.[19]

HCC cases

All patients aged 65 years and older diagnosed with HCC in SEER registries who were also enrolled in Medicare between 1994 and 1999 were eligible for inclusion. Only patients with diagnostic confirmation of HCC (ICD-O histology code = 8170) were included. Diagnostic confirmation was defined as having positive histology, cytology, laboratory test/marker study, direct visualisation, or positive radiology tests. Patients with a clinical diagnosis only or unknown method of confirmation were excluded. Patients whose HCC diagnoses were reported exclusively by death certificates or at autopsy were also excluded. For these patients, no other information regarding their HCC diagnosis was available from the medical record.[20] In addition, we excluded patients diagnosed with stomach, colon, lung, pancreatic, breast, or rectal cancers within the five years prior to the date of HCC diagnosis to further ensure the inclusion of only HCC, rather than metastatic liver cancers.

Population based non-cancer controls

Population based non-cancer controls age 65 years and older were selected from the 5% random sample of Medicare enrolled beneficiaries. Non-cancer controls were frequency matched on time of diagnosis in a 3:1 ratio to patients with HCC. Random selection without replacement was used to ensure that no non-cancer control was assigned to more than one patient with HCC.

Risk factor information
Definitions of risk factors

We identified diabetes (ICD-9 code: 250), HBV (ICD-9 codes: 070.22, 070.23, 070.32, 070.33, V02.61), haemochromatosis (ICD-9 code: 275.0), HCV, and alcoholic liver disease.[21] HCV was defined by the presence of ICD-9 codes for HCV (ICD-9 codes: 070.41, 070.44, 070.51, 070.54, V02.62) or unspecified hepatitis (ICD-9 codes: 070.9, 571.4, 571.8, 571.9). Prior to 1992, a HCV specific ICD-9 code was not available; all patients with HCV were classified as having unspecified hepatitis. Alcoholic liver disease was defined by the presence of ICD-9 codes for alcoholic fatty liver disease (ICD-9 code: 571.0), alcoholic hepatitis (ICD-9 code: 571.1), alcoholic cirrhosis of the liver (ICD-9 code: 571.2), alcoholic liver damage (ICD-9 code: 571.3), and cirrhosis (ICD-9 codes: 571.5, 571.6) in the presence of alcoholism (ICD-9 code: 291, 303, 305.0). Patients with HCC with none of these risk factors (HCV, HBV, alcoholic liver disease, haemochromatosis) were denoted as idiopathic. We also examined other potential risk factors including obesity (ICD-9 codes: 278.0, 259.9), dyslipidaemia (ICD-9 codes: 272.0, 272.1, 272.2, 272.3, 272.4), human immunodeficiency virus (HIV) (ICD-9 codes: 042–044, V08), and non-specific cirrhosis. Non-specific cirrhosis was defined as the presence of cirrhosis (ICD-9 codes: 571.5, 571.6) without the presence of HCV, HBV, alcoholic liver disease, or haemochromatosis.

Risk factors were identified based on Medicare part A or B claims for the three years preceding and two years succeeding the date of HCC diagnosis or until death, with the exception of diabetes, obesity, and dyslipidaemia. HCC cases and non-cancer controls provided information for an equal number of years for all risk factors. This time span was selected based on a sensitivity analysis that demonstrated only a small increase in the ability to identify more recorded risk factors among HCC cases and non-cancer controls beyond this time period. However, diabetes can potentially be caused by end stage liver disease and therefore we did not consider a diagnosis of diabetes that occurred after the index date of HCC diagnosis. A diagnosis of obesity and dyslipidaemia that occurred after their index date of diagnosis were also excluded because of their association with diabetes. We assumed that all other risk factors (HCV, HBV, alcoholic liver disease, haemochromatosis, non-specific cirrhosis, and HIV) that were recorded after the cancer diagnosis were acquired prior to the cancer diagnosis. These risk factors were unlikely to be acquired late in life or to be caused by HCC.

Demographic information

We collected information on age (<75, ≥75) sex (men, women), race (White, Black, Hispanic, Asian, other race), geographical region (Utah, Atlanta, Connecticut, Detroit, Hawaii, Iowa, Los Angeles, New Mexico, San Francisco, San Jose, and Seattle) and Medicare/Medicaid dual enrolment. The ''state buy-in'' variable in Medicare indicates whether a

Table 1 Comparison of demographic characteristics and risk factors among Medicare enrolled patients diagnosed with hepatocellular carcinoma (HCC) and randomly assigned non-cancer controls (1:3 ratio) between 1994 and 1999

	HCC cases (n = 2061)		Non-cancer controls (n = 6183)		Unadjusted OR (95% CI)
	n	%	n	%	
Age					
<75 y	922	44.7	2923	47.3	Reference
≥75 y	1139	55.3	3260	52.7	1.11 (1.01–1.22)
Sex					
Female	709	34.4	3935	63.6	Reference
Male	1352	65.6	2248	36.4	3.34 (3.01–3.71)
Race					
White	1367	66.3	5080	82.1	Reference
Black	181	8.8	418	6.8	1.61 (1.33–1.92)
Hispanic	91	4.4	163	2.6	2.08 (1.59–2.70)
Asian	267	13.0	301	4.9	3.30 (2.77–3.93)
Other	155	7.5	221	3.6	2.63 (2.08–3.23)
SEER registry					
Atlanta	102	5.0	368	6.0	Reference
Utah	54	2.6	305	4.9	0.64 (0.44–0.92)
Connecticut	222	10.8	804	13.0	0.99 (0.76–1.30)
Detroit	329	16.0	935	15.1	1.27 (0.99–1.63)
Hawaii	100	4.9	180	2.9	2.00 (1.44–2.78)
Iowa	176	8.5	835	13.5	0.76 (0.58–0.99)
Los Angeles	479	23.2	1027	16.6	1.68 (1.32–2.15)
New Mexico	108	5.2	313	5.1	1.24 (0.91–1.70)
San Francisco	223	10.8	515	8.3	1.56 (1.19–2.05)
San Jose	93	4.5	300	4.9	1.12 (0.81–1.54)
Seattle	175	8.5	601	9.7	1.05 (0.80–1.39)
Medicare/Medicaid dual enrolment	559	27.1	894	14.5	2.20 (1.95–2.48)
Risk factors					
Diabetes	892	43.3	1198	19.4	3.18 (2.85–3.54)
HCV	406	17.6	80	1.3	18.71 (14.64–23.93)
HBV	182	8.8	14	0.2	42.68 (24.72–73.68)
Alcoholic liver disease	453	22.0	22	0.4	78.89 (51.25–121.45)
Haemochromatosis	68	3.3	20	0.3	10.51 (6.37–17.35)
HIV	7	0.3	5	0.1	4.21 (1.34–13.28)

SEER, Surveillance, Epidemiology, and End-Results Program; HBV, hepatitis B virus; HCV, hepatitis C virus; HIV, human immunodeficiency virus; OR, odds ratio; 95% CI, 95% confidence interval.

third party payer, usually Medicaid, was paying for a beneficiary's Medicare part B premiums, and served as a proxy for lower socioeconomic status. These individuals were considered Medicare/Medicaid dually enrolled.

Statistical analysis

We compared the demographic features and prevalence of major risk factors between patients with HCC and non-cancer controls. Among patients with HCC, the proportions of patients diagnosed by testing modality obtained from SEER were calculated. We also examined the risk of HCC associated with several other features of metabolic syndrome, including hypertension, congestive heart failure, coronary atherosclerosis, ischaemic heart disease, and peripheral vascular disease. In addition, we compared demographic features and prevalence of risk factors associated with diabetes between HCC patients with a major known risk factor (HCV, HBV, alcoholic liver disease, haemochromatosis) with those with no risk factors (idiopathic). χ^2 tests were conducted for categorical variables and t tests for continuous variables.[22] Unadjusted odds ratios (OR) and 95% confidence interval (CI), as well as p values, were calculated for each variable. A sensitivity analysis was performed in which diabetes was defined based on diagnoses made between two and three years prior to the date of HCC diagnosis.

Five separate multiple logistic regression models were constructed to examine risk of HCC (v non-cancer controls) with diabetes, HCV, HBV, alcoholic liver disease, and haemochromatosis.[23] We tested for interactions between diabetes and each of these risk factors. In addition, a model restricted to patients without HCV, HBV, alcoholic liver disease, and haemochromatosis was examined. All models were adjusted for age, sex, race, SEER registry, and Medicare/Medicaid dual enrolment. Wald χ^2 tests were used to determine the significance of each variable. Adjusted ORs and 95% CIs were calculated for each parameter estimate.

RESULTS

We identified 3001 patients aged 65 years and older in the SEER-Medicare database with HCC between 1994 and 1999. Of these cases, 2061 patients with HCC satisfied our criteria for inclusion in the study cohort. Patients were excluded for enrolment in a Medicare HMO plan during the three years prior to or following the date of HCC diagnosis (n = 742; 24.7%), enrolment in Medicare part A and part B for less than three years prior to the index date (n = 114; 3.8%), diagnosis of stomach, colon, lung, pancreatic, breast, or rectal cancer within the five year period before the date of HCC diagnosis (n = 58; 1.9%), and HCC reported only from autopsy or death certificate (n = 26; 0.9%).

We identified 127 466 non-cancer controls aged 65 years and older who were enrolled in Medicare between 1994 and 1999. The 36 632 non-cancer controls were excluded for the following reasons: enrolment in a Medicare HMO plan during the three years prior to randomly assigned index date (n = 32 346; 25.4%) and enrolment in Medicare part A and part B for less than three years prior to the randomly assigned index date (n = 4286; 3.4%). Among the remaining eligible non-cancer control group (n = 92 120), a total of 6183

Table 2 Results from separate multiple logistic regression analyses examining the risk of HCC (v non-cancer controls) with each of diabetes, HCV, HBV, alcoholic liver disease, and haemochromatosis

	Adjusted OR	95% CI	p Value
Diabetes	3.08	2.74–3.46	<0.0001
Excluding HCV (n=7758)	2.99	2.64–3.40	<0.0001
Excluding HBV (n=8048)	3.11	2.76–3.50	<0.0001
Excluding alcoholic liver disease (n=7769)	2.94	2.59–3.34	<0.0001
Excluding haemochromatosis (n=8156)	3.08	2.73–3.47	<0.0001
Excluding patients without any of the above risk factors (n=7042)	2.87	2.49–3.30	<0.0001
HCV without diabetes	24.42	17.49–34.11	<0.0001
HCV and diabetes	36.88	2.64–3.40	<0.0001
HBV	23.94	13.65–41.99	<0.0001
Alcohol liver disease	69.62	44.92–107.90	<0.0001
Haemochromatosis	8.88	5.24–15.07	<0.0001

HCC, hepatocellular carcinoma; HBV, hepatitis B virus; HCV, hepatitis C virus; OR, odds ratio; 95% CI, 95% confidence interval.
All models adjusted for age, sex, race, SEER registry, and Medicare/Medicaid dual enrolment (n=8244).
Interaction terms between HBV and diabetes, alcoholic liver disease and diabetes, and haemochromatosis and diabetes were not significant at a p value of <0.05.

non-cancer controls were frequency matched on the time of diagnosis in a 3:1 ratio to patients with HCC.

Demographic features

Mean age in patients with HCC (76.1 (SD 5.9) years) and non-cancer controls (76.4 (SD 6.9) years) was not significantly different. However, patients with HCC were significantly less likely to be female and of White race than non-cancer controls. Several significant differences by SEER registry were also observed between cases and controls (table 1).

Among all testing modalities (histology, cytology, microscopic confirmation (method not specified), laboratory test/marker study, direct visualisation, or positive radiology tests), the majority of patients with HCC were diagnosed by histology (n = 1279; 62.1%), followed by cytology (n = 401; 19.4%), radiology (n = 314; 15.2%), laboratory test/marker study (n = 55; 2.7%), microscopic confirmation (method not specified) (n = 10; 0.5%), and direct visualisation (n = 2; 0.1%). These categories are constructed by SEER in a hierarchical mutually exclusive fashion in the order shown above. For example, the presence of positive histological examination supersedes all other modalities, while positive radiology diagnosis is considered only in the absence of all diagnostic modalities.

Diabetes as a risk factor for HCC

We found that 43.3% of HCC patients and 19.4% of non-cancer controls had diabetes, as defined by diagnoses recorded during the three years preceding the date of HCC

Table 3 Multiple logistic regression analysis examining association between hepatocellular carcinoma (HCC) and diabetes, adjusting for age, sex, race, SEER registry, and Medicare/Medicaid dual enrolment (n=8244)

	OR	95% CI	Wald χ^2 test	p Value
Diabetes				
No	1.00	–	–	Reference
Yes	3.08	2.74–3.46	469.04	<0.0001
Age at HCC diagnosis				
<75 y	1.00	–	–	Reference
≥75 y	0.69	0.62–0.77	40.68	<0.0001
Sex				
Female	1.00	–	–	Reference
Male	3.70	3.23–4.17	500.99	<0.0001
Race				
White	1.00	–	–	Reference
Black	1.27	1.03–1.57	4.88	0.0272
Hispanic	1.22	0.90–1.65	1.57	0.2096
Asian	2.37	1.89–2.98	55.17	<0.0001
Other	2.14	1.67–2.73	36.72	<0.0001
Registry				
Atlanta	1.00	–	–	Reference
Utah	0.62	0.42–0.92	5.55	0.0185
Connecticut	0.89	0.66–1.20	0.59	0.4419
Detroit	1.07	0.81–1.43	0.25	0.6204
Hawaii	0.67	0.45–0.99	4.03	0.0446
Iowa	0.86	0.63–1.16	0.9928	0.3191
Los Angeles	0.68	0.51–0.88	7.86	0.0050
New Mexico	0.97	0.69–1.38	0.0230	0.8794
San Francisco	0.71	0.53–0.96	4.88	0.0272
San Jose	0.65	0.45–0.93	5.49	0.0191
Seattle	0.89	0.66–1.21	0.5351	0.4645
Medicare/Medicaid dual enrolment	2.19	1.87–2.56	96.36	<0.0001

OR, odds ratio; 95% CI, 95% confidence interval.

Table 4 Comparison of demographic characteristics and risk factors among hepatocellular carcinoma (HCC) patients with idiopathic disease with those with known risk factors (n = 2061)

	Frequency of HCC patients with idiopathic disease (%) (n = 1229)	Frequency of HCC patients with known risk factors (%) (n = 832)	Unadjusted OR (95% CI)	p Value
Age at HCC diagnosis				
<75 y	473 (38.5)	449 (54.0)	–	Reference
≥75 y	756 (61.5)	383 (46.0)	1.87 (1.57–2.24)	<0.0001
Sex				
Female	448 (36.5)	261 (31.4)	–	Reference
Male	781 (63.5)	571 (68.6)	0.80 (0.66–0.96)	0.0172
Race				
White	858 (69.8)	509 (61.2)	–	Reference
Black	121 (9.9)	60 (7.2)	1.20 (0.86–1.66)	0.2840
Hispanic	48 (3.9)	43 (5.2)	0.66 (0.43–1.01)	0.0564
Asian	117 (9.5)	150 (18.0)	0.46 (0.35–0.60)	<0.0001
Other	85 (6.9)	70 (8.4)	0.72 (0.51–1.01)	0.0541
Registry				
Atlanta (%)	63 (61.8)	39 (38.2)	1.09 (0.73–1.65)	0.6524
Utah (%)	40 (74.1)	14 (25.9)	1.97 (1.06–3.64)	0.0284
Connecticut (%)	142 (64.0)	80 (36.0)	1.23 (0.92–1.64)	0.1636
Detroit (%)	209 (63.5)	120 (36.5)	1.22 (0.95–1.55)	0.1163
Hawaii (%)	58 (58.0)	42 (42.0)	0.93 (0.62–1.40)	0.7332
Iowa (%)	121 (68.7)	55 (31.3)	1.54 (1.11–2.15)	0.0099
Los Angeles (%)	239 (49.9)	240 (50.1)	0.60 (0.48–0.73)	<0.0001
New Mexico (%)	67 (62.0)	41 (38.0)	1.11 (0.75–1.66)	0.6006
San Francisco (%)	99 (44.4)	124 (55.6)	0.50 (0.38–0.66)	<0.0001
San Jose (%)	48 (48.4)	45 (51.6)	0.71 (0.47–1.08)	0.1068
Seattle (%)	118 (67.4)	57 (32.6)	1.44 (1.04–2.01)	0.0280
Medicare/Medicaid dual enrolment	301 (24.5)	258 (31.0)	0.72 (0.59–0.88)	0.0011
Risk factors				
Diabetes	504 (41.0)	388 (46.6)	0.80 (0.67–0.95)	0.0114
Obesity	88 (7.2)	58 (7.0)	1.02 (0.73–1.45)	0.8696
Dyslipidaemia	400 (32.5)	287 (34.5)	0.91 (0.76–1.10)	0.3572
Non-specific cirrhosis	285 (23.2)	72 (8.7)	3.19 (2.42–4.20)	<0.0001

diagnosis. When the definition of diabetes excluded diagnoses made only within one year or two years preceding the date of HCC diagnosis, the proportions of diabetes were 37.9% and 31.6% for HCC patients and 16.0% and 12.1% for non-cancer controls. This analysis indicated that the significant increase in the prevalence of diabetes among patients with HCC and non-cancer controls is less likely to be a result of HCC.

In the unadjusted analysis, diabetes was associated with a threefold increase in the odds of HCC (table 1). Among HCC patients with diabetes, 56.5% did not have any specific major risk factors while 22.3% had HCV, 8.2% HBV, 24.2% alcoholic liver disease, and 4.7% haemochromatosis (patients could have multiple risk factors).

In the multivariate analysis, adjusting for age, sex, race, geographic region, and Medicare/Medicaid dual enrolment, the odds of HCC patients being diagnosed with diabetes increased by almost threefold (tables 2, 3). The effect of diabetes was tested after excluding patients with each risk factor, as well as all patients with any of these risk factors. The increase in the odds for HCC patients with diabetes remained when the analysis was restricted to patients without HCV (OR 2.99 (95% CI 2.64–3.40)), HBV (OR 3.11 (95% CI 2.76–3.50)), alcoholic liver disease (OR 2.94 (95% CI 2.59–3.34)), haemochromatosis (OR 3.08 (95% CI 2.73–3.47)), or any of these risk factors (OR 2.87 (95% CI 2.49–3.30)). A significant positive statistical interaction was found between HCV and diabetes (p<0.0001). The combined presence of HCV and diabetes was associated with a 37-fold increase in HCC. There were no other significant interactions among diabetes and other risk factors.

In these models, an increase in the odds of having HCC was also associated with younger age, male sex, non-White race, and having Medicare/Medicaid dual enrolment (table 3). Significant positive interactions were found between diabetes and age, sex, and race (p<0.0001) where the combined presence of diabetes with each of older age, male sex, and non-Caucasian race further increased the risk of HCC.

Other potential HCC risk factors
As expected, the proportions of HCC patients with HCV, HBV, alcoholic liver disease, and haemochromatosis were significantly greater compared with non-cancer controls (table 1). Among patients with HCC, 43.3% had diabetes, 17.6% HCV, 8.8% HBV, 22.0% alcoholic liver disease, 3.3% haemochromatosis, and 0.3% HIV. Among non-cancer controls, 19.4% had diabetes, 1.3% HCV, 0.2% HBV, 0.4% alcoholic liver disease, 0.3% haemochromatosis, and 0.1% HIV.

In separate multiple logistic regression models examining the association between risk of HCC and each of these risk factors while adjusting for age, sex, race, geographical region, and Medicare/Medicaid dual enrolment, patients with HCC continued to be associated at a 24, 24, 70, and 9-fold increased odds of HCV, HBV, alcoholic liver disease, and haemochromatosis, respectively, compared with non-cancer controls (table 2).

Other features of metabolic syndrome as risk factors for HCC
The proportions of HCC patients with hypertension, congestive heart failure, coronary atherosclerosis, ischaemic heart disease, and peripheral vascular disease were significantly greater compared with non-cancer controls. Among patients with HCC, 67.7% had hypertension, 38.9% congestive heart failure, 25.6% coronary atherosclerosis, 20.3% ischaemic heart disease, and 15.4% peripheral vascular

disease. Among non-cancer controls, 49.2% had hypertension, 18.5% congestive heart failure, 16.1% coronary atherosclerosis, 11.5% ischaemic heart disease, and 9.8% peripheral vascular disease. In univariate analyses examining the association between risk of HCC and each of these features of metabolic syndromes, congestive heart failure was associated with the greatest risk of HCC (OR 2.80 (95% CI 2.56–3.07)) followed by hypertension (OR 2.16 (95% CI 2.16–1.97)), ischaemic heart disease (OR 1.96 (95% CI 1.77–2.20)), coronary atherosclerosis (OR 1.79 (95% CI 1.62–1.98)), and peripheral vascular disease (OR 1.67 (95% CI 1.48–1.89)).

Comparison of patients with a known risk factor with those with idiopathic HCC

Patients with idiopathic HCC were older and more likely to be of White race but no significant gender differences were observed (table 4). Idiopathic HCC constituted the majority of patients in Utah (74.1%), Iowa (68.7%), Seattle (67.4%), Detroit (63.5%), and New Mexico (62.0%). A lower proportion of HCC patients with idiopathic disease were dually enrolled in Medicare/Medicaid programmes compared with patients with a known risk factor (24.5% v 31.0%). The proportion of idiopathic HCC patients with non-specific cirrhosis was significantly greater compared with HCC patients with a known risk factor but the proportion of idiopathic patients with diabetes was significant lower compared with HCC patients with a known risk factor (table 4). No significant differences were observed between these groups for obesity or dyslipidaemia.

DISCUSSION

This is the first population based study to examine the association between diabetes and HCC in the USA. We found diabetes to be an independent risk factor for HCC, regardless of the presence of HCV, HBV, alcoholic liver disease, or non-specific cirrhosis. In this study, diabetes was associated with a 2–3-fold increase in the risk of HCC. Approximately 60% of patients with HCC in this study did not have a recorded diagnosis of HCV, HBV, alcoholic liver disease, or haemochromatosis; among these patients, 47% had diabetes, which was even higher than those with other risk factors (41%). This finding suggests that diabetes may account for a significant proportion of patients with idiopathic HCC.

These results corroborate and extend findings from several non-population based as well as non-US population based studies that have identified diabetes as a risk factor for HCC. One population based study using data obtained from the Denmark cancer registry detected an increased risk of HCC among patients with diabetes alone, as well as in the presence of hepatitis, cirrhosis, alcohol related conditions, and other liver disorders.[13] Similarly, a study using the Swedish inpatient registry also found a threefold risk of liver cancer among patients hospitalised with diabetes, and a fourfold risk in the presence of hepatitis, cirrhosis, and alcoholism.[15]

We also found that diabetes in the presence of HCV, HBV, alcoholic liver disease, and haemochromatosis increased the risk of developing HCC beyond the risk associated with diabetes in the absence of these risk factors. Although a positive interaction between diabetes and these risk factors could exist where the combination of risk factors act synergistically to increase the risk of HCC, this finding represents an increased risk of diabetes with these factors either directly (for example, HCV) or through underlying liver disease. In this study, the statistical evidence of significant interaction was found only between diabetes and HCV. Two previous non population based studies found that diabetes increased the risk of liver cancer only in the presence of hepatitis or cirrhosis.[12 16]

In this study, HCV, HBV, alcoholic liver disease, and haemochromatosis accounted for only approximately 40% of HCC. This finding may indicate a limitation of the data source but is similar to findings from previous studies reporting that 20–50% of HCC patients in the USA have idiopathic disease.[3 4 24] Recent studies have suggested that diabetes and/or NAFLD could account for at least a portion of these ''idiopathic'' cases.[12 15] Findings from this study support this hypothesis; the presence of diabetes alone accounted for approximately 32% of patients with idiopathic HCC.

Diabetes, as part of the insulin resistance syndrome, has been implicated as a risk factor for NAFLD, including its most severe form NASH. NASH has been identified as a cause of both cryptogenic cirrhosis and HCC.[5 6] One would expect that patients at risk of developing NASH related HCC are more likely to be older, female, of White race, and to have other features of insulin resistance, such as obesity and dyslipidaemia. We observed a greater proportion of older as well as White patients with idiopathic HCC. We also observed a significantly greater proportion of idiopathic HCC cases in geographical areas with predominately White populations (for example, Utah (74% v 26%) and Iowa (69% v 31%)) compared with the remaining SEER regions. Lastly, we also observed fewer idiopathic HCC patients with lower socio-economic levels who are at higher risk of viral hepatitis and alcoholic liver disease. On the other hand, we found no significant association between idiopathic HCC and sex, obesity, or dyslipidaemia.

Limitations of this study are related to the completeness and accuracy of the information on risk factors in the SEER-Medicare linked database. To maximise the possibility of capturing risk factor information from claims data, we searched for risk factor information for up to a five year period (three years before and up to two years after HCC diagnosis) for HCV, HBV, alcoholic liver disease, and haemochromatosis, and up to a three year time period prior to HCC diagnosis for diabetes. The choice of this five year time span was selected based on our sensitivity analysis that demonstrated only a small increase in the ability to identify more recorded risk factors among HCC cases and non-cancer controls beyond this time period. For our analysis sample of patients 65 years or older with continuous Medicare part A and part B enrolment and no enrolment in an HMO plan, Medicare files have been reported to capture 100% of their Medicare claims for tests, procedures, outpatient visits, and hospitalisation.[19]

Given the case control design and relatively limited time frame for risk factor information, we could not investigate the duration-response relationship between diabetes and HCC. One study found that patients with a longer duration since the onset of diabetes had a greater risk of developing HCC compared with patients more recently diagnosed with diabetes.[12] We also cannot be completely certain that diabetes in HCC patients did not develop as a consequence of underlying liver disease.[25] However, we took several steps to limit and adjust for this possibility by considering only diabetes related risk factor information prior to the date of cancer diagnosis. In addition, we examined the prevalence of diabetes excluding the one and two years prior to HCC diagnosis and found no significant differences in the prevalence of diabetes among patients with HCC.

Strengths of this study are related to its data source. Firstly, the SEER registries are selected to represent the entire US population, and therefore our overall findings should be generalisable to the entire US population aged 65 years and older.[19] Secondly, we had complete ascertainment of patients with HCC, which also facilitates the generalisability of our results. The SEER programme maintains at least a 98% completeness rate for case ascertainment. Thirdly, all cases of

HCC included in this analysis were confirmed by pathology, radiology, and/or laboratory testing. Fourthly, we selected our cohort to obtain complete Medicare claims data for all patients with HCC and non-cancer controls thus minimising the possibility of losing any recorded encounters with the health care system. Lastly, there was no potential for recall bias between cases and controls as all risk factor information was based on diagnostic codes contained in the Medicare claims files.

In conclusion, our results indicate that diabetes is associated with an increased risk of HCC among persons 65 years and older, regardless of the presence of HCV, HBV, alcoholic liver disease, haemochromatosis, or other underlying liver disease. Our findings also suggest that HCV could potentially interact with diabetes to further increase the risk of developing HCC. Further studies are needed to clarify the role of diabetes on the progression of pre-existing major risk factors, and how the presence of multiple risk factors with diabetes affect the development of HCC.

. . . .

Authors' affiliations

J A Davila, R O Morgan, Section of Health Services Research, Houston Veterans Affairs Medical Center and Baylor College of Medicine, Houston, Texas, USA

Y Shaib, H B El-Serag, Section of Health Services Research, and Section of Gastroenterology, Houston Veterans Affairs Medical Center and Baylor College of Medicine, Houston, Texas, USA

K A McGlynn, Division of Cancer Epidemiology and Genetics, NCI/DHHS, USA

REFERENCES

1 **El-Serag HB**, Mason AC. Rising incidence of hepatocellular carcinoma in the United States. N Engl J Med 1999;**340**:745–50.
2 **El-Serag HB**, Davila J, Petersen N, et al. The continuing rise in hepatocellular carcinoma in the United States: An Update. Ann Intern Med, 2003 18, **139**:817–23.
3 **El-Serag HB**, Mason AC. Risk factors for the rising rates of primary liver cancer in the United States. Arch Intern Med 2000;**160**:3227–30.
4 **Hassan MM**, Frome A, Patt YZ, et al. Rising prevalence of hepatitis C virus infection among patients recently diagnosed with hepatocellular carcinoma in the United States. J Clin Gastroenterol 2002;**35**:266–9.
5 **Bugianesi E**, Leone N, Vanni E, et al. Expanding the natural history of nonalcoholic steatohepatitis: From cryptogenic cirrhosis to hepatocellular carcinoma. Gastroenterology 2002;**123**:134–44.
6 **Caldwell SH**, Oelsner DH, Iezzoni JC, et al. Cryptogenic cirrhosis: Clinical characterization and risk factors for underlying disease. Hepatology 1999;**29**:664–9.
7 **Powell EE**, Cooksley WGE, Hanson R, et al. The natural history of nonalcoholic steatohepatitis: A follow-up study of forty-two patients for up to 21 years. Hepatology 1989;**11**:74–80.
8 **Lee RG**. Nonalcoholic steatohepatitis: A study of 49 patients. Hum Pathol 1989;**20**:594–8.
9 **Dixon JB**, Bhathal PS, O'Brien PE. Nonalcoholic fatty liver disease: Predictors of nonalcoholic steatohepatitis and liver fibrosis in the severely obese. Gastroenterology 2001;**121**:91–100.
10 **Angulo P**, Keach JC, Batts KP, et al. Independent predictors of liver fibrosis in patients with nonalcoholic steatohepatitis. Hepatology 1999;**30**:1356–62.
11 **Petrides AS**, Vogt C, Schulze-Berge D, et al. Pathogenesis of glucose intolerance and diabetes mellitus in cirrhosis. Hepatology 1994;**19**:616–27.
12 **El-Serag HB**, Tran T, Everhart J. Diabetes increases the risk of chronic liver disease and hepatocellular carcinoma: a cohort study among US veterans. Gastroenterology 2003;**126**:460–8.
13 **Wideroff L**, Gridley G, Mellemkjaer L, et al. Cancer incidence in a population-based cohort of patients hospitalized with diabetes mellitus in Denmark. J Natl Cancer Inst 1997;**89**:1360–5.
14 **Adami HO**, McLaughlin J, Ekbom A, et al. Cancer risk in patients with diabetes mellitus. Cancer Causes Control 1991;**2**:307–14.
15 **Adami HO**, Chow WH, Nyren O, et al. Excess risk of primary liver cancer in patients with diabetes mellitus. J Natl Cancer Inst 1996;**88**:1472–7.
16 **Fugino Y**, Mizoue T, Tokui N, et al. Prospective study of diabetes mellitus and liver cancer in Japan. Diabetes Metab Res Rev 2001;**17**:374–9.
17 **Yu MC**, Tong MJ, Govindarajan S, et al. Nonviral risk factors for hepatocellular carcinoma in a low-risk population, the non-Asians of Los Angeles County, California. J Natl Cancer Inst 1991;**83**:1820–6.
18 **Hassan MM**, Hwang LY, Hatten CJ, et al. Risk factors for hepatocellular carcinoma: synergism of alcohol with viral hepatitis and diabetes mellitus. Hepatology 2002;**36**:1206–13.
19 **Warren JL**, Klabunde CN, Schrag D, et al. Overview of the SEER-Medicare data: content, research applications, and generalizability to the United States elderly population. Med Care 2002;**40**(8 Suppl):IV-3–18.
20 The SEER Program Code Manual, 3rd ed. Bethesda: National Cancer Institute, 1998.
21 **International Classification of Diseases, Ninth Revision**. In:Clinical Modification, 4th edn., Washington, DC: Public Health Service, US Dept of Health and Human Services 1991.
22 **Fleiss JL**, Levin B, Paik MC. Statistical methods for rates and proportions, 3rd edn., Hoboken: John Wiley and Sons, Inc. 2003:250–4.
23 **Hosmer DW Jr**, Lemeshow S. Applied logistic regression. New York: John Wiley and Sons, Inc, 1989:106–18.
24 **Kulkerni K**, Barcak E, El-Serag HB, et al. Increasing incidence of hepatocellular carcinoma among US born residents of Houston, TX. Am J Gastroenterol 2003;**98**:S84.
25 **Perseghin G**, Mazzaferro V, Sereni LP, et al. Contribution of reduced insulin sensitivity and secretion to the pathogenesis of hepatogenous diabetes: effect of liver transplantation. Hepatology 2000;**31**:694–703.